



Cornell University Library

Report of the Activities of the Steering Committee of the Metadata Working Group 2001-

2004-06-28

Bill Kehoe
Martin Kurth
David Ruddy
Nathan Rupp
Kari Smith
Kizer Walker
Elaine Westbrooks, Chair



TABLE OF CONTENTS

EXECUTIVE SUMMARY.....	1
INTRODUCTION.....	2
COMMITTEE.....	3
FORUMS	
Announcements.....	3
Topics.....	3-4
Speakers.....	4-5
Attendees/Audience.....	5
WEBSITE.....	6
SUGGESTIONS FOR IMPROVEMENT	
Evaluation.....	7
Role of Steering Committee.....	7
Funding.....	8
CONCLUSION.....	8
APPENDIX.....	9

EXECUTIVE SUMMARY

The Cornell University Library (CUL) Metadata Working Group (MD-WG) was formed in 2001 as an open forum for all staff to discuss metadata issues facing the library. To that end, it has held thirty forums over the last three years on a variety of metadata issues and maintains a website which keeps the CUL community informed about its activities.

Since its inception, the MD-WG steering committee, occasionally in concert with other campus organizations, has brought nearly fifty speakers from as far away as Australia and the United Kingdom to as near as Syracuse and Dryden, New York. The majority of presenters represented numerous departments from Cornell University, (Legal information Institute and Digital Library Research Group) but there have also been also speakers from outside institutions such as Columbia University, and UKOLN (formerly UK Office for Library Networking). Some of the topics have included Computational Linguistics for Metadata Building and Dublin Core structured values.

The MD-WG has earned a reputation for consistently bringing provocative, timely, and interesting speakers and topics to the CUL, and furthermore, has been instrumental in bringing metadata issues to the attention of a wider audience at CUL. Although our data are not comprehensive, our statistics suggest that the forums are attracting more people from increasingly diverse departments inside and outside of the CUL. In the past three years, approximately a total of 650 people from twelve CUL departments attended thirty forums. Of that 650 in total attendance, 100 of them came from five departments outside the CUL.

The purpose of this report is to document the progress of the MD-WG Steering Committee during my tenure as chair. In addition to highlighting a few facts about the forums (topics, speakers, attendees), this report covers the redesign of the website and lists a number of the committee's achievements between 2001 and 2004. Finally, we have identified a few suggestions and areas for improvement that could help the MD-WG forums and committee flourish.

Introduction

In August of 2001, Ross Atkinson, Associate University Librarian for Collections, established the following charge for the MD-WG:

“The Metadata Working Group will serve as an open forum for all interested CUL staff to learn about and to discuss information on all issues relating to metadata. Specifically, the purpose of the group will be:

- To exchange information about metadata and its application to all library functions; this exchange will include information about uses of metadata in CUL as well as information about broader metadata developments nationally and internationally.
- To provide an opportunity for feedback on the potential and the impact of intended developments or uses of metadata in CUL.
- To make recommendations on policy decisions relating to metadata that needs to be made by CUL.
- The agenda for the Working Group will be set by the Steering Committee, the members of which will rotate every two years.”

On December 20, 2001, six months after the MD-WG was formed, Tom Turner, metadata librarian and then chair of the steering committee, issued a summary report to Sarah Thomas, Carl A. Kroch University Librarian, which stated:

“The meetings have been very well attended, usually by 25 to 30 people per meeting, and have generated lively discussion and interest in metadata issues. Many people have commented to members of the steering committee that they really enjoy the sessions and the chance to learn about and contribute to discussions about these topics.... We would like to thank you and LMT for your encouragement in creating this forum and for the opportunity to contribute to fostering a discussion among CUL staff. It has been a rewarding experience for all of us.”

Fortunately Tom’s statement remains true two and a half years later.

THE COMMITTEE

The Steering Committee's primary responsibility is to set the agenda for the forums. Seemingly a simple task, setting the agenda requires a great deal of work. First, having members that are quite knowledgeable about the latest developments in metadata and digital library technology is a boon. However, it is not necessary that each and every member be versed in the language of metadata. Rather it is important that the committee remains to be composed of individuals representing a variety of departments, skill levels, interests, and libraries. This mixture ensures that forums will be targeted to a variety of people from various departments and libraries that also possess a variety of skills and experience. It is apparent that the committee appointments have been well thought out, which would account for the great teamwork that I have witnessed as chair. I contend that the success of the forums is a direct result of the collective knowledge, team work, and resolve to repeatedly address timely topics and identify the most appropriate speakers. The committee members that I have worked with have been highly motivated, professional, and I can not thank them enough for their service. For a list of committee members, see APPENDIX, p. 9.

THE FORUMS

Announcements

Although the Forums are typically scheduled for the third Friday of every month, an announcement is posted to <CU-Lib> two to three weeks before the forum is scheduled to occur. The announcement includes the speaker, an abstract, as well as online links to more relevant information. Many staff have stated that the announcements are quite informative and are often reviewed when one is able or unable to attend the forum. Since 2001, greater effort has been made to attract more attendees from CIT and Information Science/Computer Science, so announcements have been posted to <lsgeneral-l>, the Information Science General Listserv as well as CIT's general list.

Topics

The strength of the MD-WG has been its ability to program high quality forums on a monthly basis-- thirty to date-- that attracted a diverse group of attendees within the CUL and without. Since the main focus of the forum is metadata, many presentations have focused on either individual schemas and standards such as Dublin Core, METS, geospatial metadata, Library Application Profile, and VRA or the digital library systems such as DSpace, Uportal, ArXiv, LUNA, and the NSDL that utilize these standards. However, the majority of topics have been about issues that are tangentially relevant to metadata: preservation, resource

persistence, the semantic web, etc. Metadata appears to be more integrated into the work, digital library and otherwise, that is being done inside and outside of CUL and that is reflected in the broadening scope of the topics.

In addition to the thirty regularly scheduled forums, the MD-WG has co-sponsored three and hosted two special forums. With Karen Calhoun, AUL for Technical Services, the MD-WG has cosponsored programs featuring Sandy Hurd (Electronic Resource Management), Harry Samuels (OpenURL), and Juha Hakala (Uniform Resource Names). In addition, the MD-WG hosted a special event to honor Tom Turner and to celebrate the anniversary of the Dublin Core Metadata Initiative in March of 2003. The other special forum, held in June of 2003, consisted of a panel of international experts from the Dublin Core Usage Board: Stuart Sutton, Thomas Baker, Andy Powell, Andrew Wilson, and Diane Hillmann. The CUL community was able to learn about the latest issues being discussed by the board prior to public disclosure of such information.

Speakers

If we include the co-sponsored and special forums, a total of forty-five people have presented at thirty forums. Sixty percent of the presenters come from the CUL. Karen Calhoun, Jonathan Corson-Rikert, Bill Kehoe, Martin Kurth, and David Ruddy, are just a few speakers who have presented research.

Another twenty percent of speakers come from Cornell departments outside of the library. The research staff of Donna Bergmark, Naomi Dushay, Diane Hillmann, Carl Lagoze, and Simeon Warner from the Digital Library Research Group (Information Science/Computer Science Department) have discussed a number of innovative projects. The fact that eighty percent of the presenters are from Cornell indicates that we have a strong community of experts who can speak authoritatively on a number of subjects relevant to metadata and digital library development.

The remaining twenty percent of speakers come from institutions outside of Cornell University. The Dublin Core Usage Board panel is a case in point. The MD-WG brought a group of internationally distinguished practitioners who are known for their experience and knowledge about preservation metadata, RDF, metadata registries, application profiles, and metadata aggregation. The Usage Board, whose mission is to "ensure an orderly evolution of the metadata terms maintained by the Dublin Core Metadata Initiative," is a knowledgeable and influential metadata group.¹ In addition, the MD-WG hosted two notable

¹Dublin Core Metadata Initiative. "DCMI Usage Board Mission and Principles" Available at: <http://www.dublincore.org/usage/documents/mission/> [Accessed 09 June 2004].

speakers, also from outside of Cornell, Elizabeth Liddy, Professor in the School of Information Studies at Syracuse University and Director of its Center for Natural Language Processing as well as Simon St. Laurent, editor with O'Reilly and Associates.

Attendees/Audience

In the beginning stages of the MD-WG, statistics about forum attendance were not typically collected. However, since 2003, the committee has been more committed to collecting data about our attendees that can tell us more about our audiences.

The following observations and/or assessments can be made from the data:

1. The vast majority of attendees come from CUL Technical Services departments.
2. The majority of attendees come from departments within Olin Library
3. The majority of attendees have attended at least half or fifteen of the forums
4. Attendees are increasingly coming from departments outside of CUL, i.e. CIT, Lab of Ornithology, and the Digital Library Research Group/Computer Science department.
5. Every department from Mann, Olin, Kroch, and Uris has been represented in attendance.
6. Twelve CUL libraries are represented in the attendance.
7. We can estimate that a total of 650 people have attended forums.
8. Of the 650 in total attendance, we can estimate that 100 of them came from five departments outside the CUL.

Although our evidence is loosely empirical, I am inclined to believe that attendees are not so much interested in metadata per se. Rather, they are interested in the systems, tools, standards, and technologies that are related to metadata. In turn, this may indicate that metadata is becoming a more integrated part of the work being done at CUL, CUL technical services, RMC, and information technology in general; as well as the work being done in information science/computer science, CIT, and the Laboratory of Ornithology. Perhaps this metadata integration "phenomenon" accounts for the diversity of attendees of the MD-WG and their interest in topics (preservation, resource persistence, the semantic web) tangentially relevant to metadata. On the other hand, this diversity may be a result of the forums being more broadly publicized. Similarly, the fact that more speakers happen to be from non-CUL departments (Digital Library Research Group) probably encourages attendance of colleagues from the same department.

THE WEBSITE <<http://metadata-wg.mannlib.cornell.edu/>>

The website has been the communications organ of the steering committee since it was first created. The website is divided into four main categories, Projects, Bibliography, Forums, and About MD-WG. Most of the information is contained in the "Forum" webpages: announcements, minutes, PowerPoint presentations, and online resources are "archived" and maintained by a sub-committee of the Steering committee, the Technical Advisory Group (Keith Jenkins, Nathan Rupp, Kari Smith, and Elaine Westbrooks). The projects webpage documents the projects undertaken by the steering committee. These projects include: maintaining a metadata booth at the Learning and Teaching with Technology Expo, shepherding the process to get the Festschrift in Honor of Tom Turner published, and selecting metadata elements for still images for the CUL Digitization Service.

In May of 2004, the Metadata Working Homepage was redesigned using XHTML, CSS, and PHP to more efficiently manage the large amount of information that had become part of the website. In addition, the changes were designed to make the website visually appealing, user-friendly, searchable, and easier to maintain. Although the main page's structure did not drastically change, the color scheme, footer, and technology behind the website changed considerably. Thanks to the use of PHP and cascading style sheets, the maintenance of the website is easier because we no longer need to change content on multiple webpages; instead the content only needs to be changed once. Another great addition to the website is the new MD-WG Logo, designed by Keith Jenkins (see report cover).

Finally, we intend to add software that can compile and intelligently present the data found in the server log so that we can learn more about how the website is being accessed and used. Referrer data should also tell us a lot about the way users encounter the website. By using the `a:link` command, we found that there are nearly a dozen libraries, syllabi, and individuals who link to the website.

SUGGESTIONS FOR IMPROVEMENT

Evaluation

Although the committee is proud of its achievements, there is room for improvement. In the next year, the MD-WG may post an online survey or tool to query forum attendees. The results of this survey can be used to document recommendations and to identify areas that need improvement. We want feedback and we want to find answers to the following questions:

- What topics should the MD-WG bring to the forum?
- Is the format and length of the forum sufficient?
- What do attendees like most and least about MD-WG Forums?
- Who is benefiting most from the forums?
- Who is our audience?
- Who should be the target audience?
- How can forums in general be improved?

The Role of the Steering Committee

In concert with the charge of the MD-WG, it has been suggested that the steering committee can play a more prominent role in:

1. Keeping abreast of metadata standards, issues, and technologies appropriate for projects at CUL;
2. Documenting the CUL's position or standing in emerging metadata matters i.e. metadata registry; and
3. Presenting and/or articulating the CUL's position to appropriate parties to ensure that the CUL is given proper representation (via vote or opinion) in standards groups such as the National Information Standards Organization (NISO), DCMI, or OCLC.

Because of its independence and neutrality within CUL, the Steering Committee is well positioned to identify problem areas that relate to metadata. More importantly, it could also take a leading role in the resolution of such problems. On the other hand, the committee has struggled with a few of the responsibilities articulated in the charge. Setting the agenda and organizing the forums is a substantial task, in and of itself. In my tenure as chair, the committee was good at identifying the potential problems that could be addressed, but it never got the opportunity to resolve any of those problems.

Funding

If the steering committee is to continue bringing local and outside experts to the CUL, it would be helpful if the committee had access to a pool of money that could be used to compensate non-CUL presenters or at least defray the costs incurred for presentation or travel. In the past, Janet McCue, Director of Mann Library and AUL for Life Sciences and Karen Calhoun, have been quite generous in providing monies on an ad hoc basis. Although this method has been effective, it is doubtful that it is sustainable or an efficient way to continue to attract experts from outside of Cornell as the committee has done the past.

CONCLUSION

As chair, I have often received compliments from numerous CUL staff that suggests that the MD-WG steering committee has been exemplary in its leadership and quality programming with regard to the education of an increasing number and variety of library staff on issues related to metadata. Utilizing the knowledge of both the CUL's in-house experts and national and international speakers, the MD-WG has offered over two dozen presentations on a broad range of important metadata-related topics. Although a particular forum may not be directly relevant to an attendee's work, he or she can be sure that the presentation will be interesting, informative, and timely. In this regard, the Committee has become an excellent example of how working groups can operate within CUL.

In the future, the steering committee faces a number of challenges in fulfilling its mission; chief among them are the following:

- Continuing to bring great programs to the CUL and,
- Obtaining designated monies and support these programs,
- Seeking feedback from forum attendees to keep the topics, presenters, and format interesting and "fresh"; and
- Realizing the responsibilities associated with the committee's charge to play a decision-making role regarding metadata in the CUL and in the larger national and international arena.

Finally, LMT should be commended for having the foresight to create a forum that provides CUL staff with an opportunity to engage in vibrant discussions on current metadata issues that might not otherwise occur in other library settings. With the continued support of LMT, the MD-WG will continue to evolve and benefit everyone in the CUL, metadata experts and non-experts alike. I have enjoyed being chair of the MD-WG Steering Committee and am confident that other committee members also found the committee service rewarding.

APPENDIX

Steering Membership

2001-2002

Adam Chandler
Martha Hsu
Jill Powell
Martin Kurth
David Ruddy
Thomas Turner (Chair)
Elaine Westbrooks

2002-2003

Martha Hsu
Martin Kurth
Jill Powell
David Ruddy
Thomas Turner (Chair)
Elaine Westbrooks (Chair)

2003-2004

Bill Kehoe
Martin Kurth
David Ruddy
Nathan Rupp
Kari Smith
Kizer Walker
Elaine Westbrooks (Chair)

2004-2005

Paul McMillan
Joy Paulson
Bill Kehoe
Nathan Rupp
Rick Silterra
Kari Smith (Chair)
Kizer Walker